



(12)发明专利

(10)授权公告号 CN 107680069 B

(45)授权公告日 2020.09.11

(21)申请号 201710766169.1

G06K 9/00(2006.01)

(22)申请日 2017.08.30

(56)对比文件

(65)同一申请的已公布的文献号

CN 104539868 A,2015.04.22

申请公布号 CN 107680069 A

CN 107004301 A,2017.08.01

(43)申请公布日 2018.02.09

US 2011/0317905 A1,2011.12.29

CN 107004296 A,2017.08.01

(73)专利权人 歌尔股份有限公司

审查员 李玥

地址 261031 山东省潍坊市高新技术产业

开发区东方路268号

(72)发明人 戴天荣 朱育革 赵大川 陈翔

(74)专利代理机构 北京市隆安律师事务所

11323

代理人 权鲜枝

(51)Int.Cl.

G06T 5/50(2006.01)

G06T 5/00(2006.01)

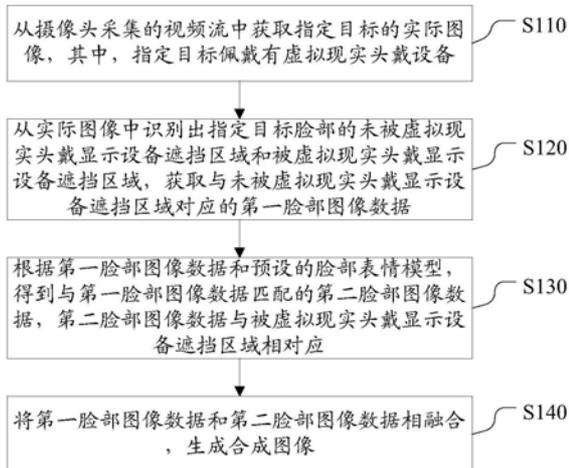
权利要求书3页 说明书7页 附图2页

(54)发明名称

一种图像处理方法、装置和终端设备

(57)摘要

本发明公开了一种图像处理方法、装置和终端设备。该方法包括：从摄像头采集的视频流中获取指定目标的实际图像；从实际图像中识别出指定目标脸部的未被虚拟现实头戴设备遮挡区域和被虚拟现实头戴设备遮挡区域，获取与未被遮挡区域对应的第一脸部图像数据；根据第一脸部图像数据和预设的脸部表情模型，得到与第一脸部图像数据匹配的第二脸部图像数据，第二脸部图像数据与被遮挡区域相对应；将第一脸部图像数据和第二脸部图像数据相融合，生成合成图像。该图像处理装置包括第一获取单元、识别单元、第二获取单元和生成单元，用于执行上述的方法步骤。本方案有利于社交双方及时获得对方的表情信息，保证社交的顺利进行，提升用户体验。



1. 一种图像处理方法,其特征在于,所述方法包括:

从摄像头采集的视频流中获取指定目标的实际图像,其中,所述指定目标佩戴有虚拟现实头戴设备;

从所述实际图像中识别出所述指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域,获取与所述未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据;

根据所述第一脸部图像数据和预设的脸部表情模型,得到与所述第一脸部图像数据匹配的第二脸部图像数据,所述第二脸部图像数据与所述被虚拟现实头戴显示设备遮挡区域相对应;所述预设的脸部表情模型是通过深度神经网络实现得到的,预设的脸部表情模型中有未被虚拟现实头戴显示设备遮挡的图像数据与被虚拟现实头戴显示设备遮挡区域的图像数据之间的关系;

将所述第一脸部图像数据和所述第二脸部图像数据相融合,生成合成图像。

2. 如权利要求1所述的图像处理方法,其特征在于,所述根据所述第一脸部图像数据和预设的脸部表情模型,得到与所述第一脸部图像数据匹配的第二脸部图像数据包括:

将所述第一脸部图像数据输入到所述预设的脸部表情模型中,以使所述脸部表情模型识别所述第一脸部图像数据,输出与所述第一脸部图像数据相匹配的第二脸部图像数据。

3. 如权利要求2所述的图像处理方法,其特征在于,所述通过深度神经网络实现得到预设的脸部表情模型包括:

获取摄像头在第一场景下采集的所述指定目标的多个第一样本图像,以及在第二场景下采集的所述指定目标的多个第二样本图像;其中,在所述第一场景下,所述指定目标佩戴有所述虚拟现实头戴设备;在所述第二场景下,所述指定目标未佩戴所述虚拟现实头戴显示设备,且各第二样本图像中包含所述指定用户的脸部状态;

从所述第一样本图像中识别出第一被遮挡区域,获取所述第一被遮挡区域信息;

根据所述第一被遮挡区域信息,对所述第二样本图像的所述指定目标脸部的与所述第一被遮挡区域对应的区域进行标记;

将所述第二样本图像中标记区域的图像放入第一指定集合中,将该第一指定集合作为深度神经网络训练时的输出集合;将所述第二样本图像中的所述指定目标脸部的未被标记区域的图像放入第二指定集合中,将该第二指定集合作为深度神经网络训练时的输入集合;所述第二指定集合与所述第一指定集合中的图像元素有一一对应的严格的输入输出对应关系;

将所述输入集合和所述输出集合中的每一对具有输入输出对应关系的图像元素输入到预设的深度神经网络中进行训练,确定未遮挡区域图像和生成的与其匹配的遮挡区域图像之间的函数关系,以使在所述第一脸部图像数据输入到所述预设的脸部表情模型时,所述预设的脸部表情模型根据输入的所述第一脸部图像数据和所述函数关系输出与其匹配的第二脸部图像数据。

4. 如权利要求3所述的图像处理方法,其特征在于,

在所述预设的深度神经网络训练过程中,所述预设的深度神经网络训练的损失函数是所述输出集合中的图像和生成的与所述输入集合中的图像相匹配的图像之间的均方差。

5. 如权利要求1所述的图像处理方法,其特征在于,所述方法还包括:

从所述实际图像中识别出被所述虚拟现实头戴设备遮挡的非脸部区域；

从所述视频流中获取所述实际图像之前的多个第三图像，从所述第三图像中提取背景图像，使用所述背景图像中与被所述虚拟现实头戴设备遮挡的非脸部区域对应的图像数据，对所述被所述虚拟现实头戴设备遮挡的非脸部区域进行去遮挡处理。

6. 如权利要求1所述的图像处理方法，其特征在于，所述方法还包括：

从所述实际图像中识别出被所述虚拟现实头戴设备遮挡的非脸部图像数据，将所述非脸部图像数据输入到预设的非脸部模型中，以使所述预设的非脸部模型识别所述非脸部图像数据，输出与被所述虚拟现实头戴设备遮挡的非脸部区域匹配的第四图像数据，根据所述第四图像数据对所述被所述虚拟现实头戴设备遮挡的非脸部区域进行去遮挡处理。

7. 一种图像处理装置，其特征在于，所述装置包括：

第一获取单元，用于从摄像头采集的视频流中获取指定目标的实际图像，其中，所述指定目标佩戴有虚拟现实头戴设备；

识别单元，用于从所述实际图像中识别出所述指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域，获取与所述未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据；

第二获取单元，用于根据所述第一脸部图像数据和预设的脸部表情模型，得到与所述第一脸部图像数据匹配的第二脸部图像数据，所述第二脸部图像数据与所述被虚拟现实头戴显示设备遮挡区域相对应；所述预设的脸部表情模型是通过深度神经网络实现得到的，预设的脸部表情模型中有未被虚拟现实头戴显示设备遮挡的图像数据与被虚拟现实头戴显示设备遮挡区域的图像数据之间的关系；

生成单元，用于将所述第一脸部图像数据和所述第二脸部图像数据相融合，生成合成图像。

8. 如权利要求7所述的图像处理装置，其特征在于，所述第二获取单元还包括：

训练模块，用于通过深度神经网络实现得到所述预设的脸部表情模型，具体用于：

获取摄像头在第一场景下采集的所述指定目标的多个第一样本图像，以及在第二场景下采集的所述指定目标的多个第二样本图像；其中，在所述第一场景下，所述指定目标佩戴有所述虚拟现实头戴设备；在所述第二场景下，所述指定目标未佩戴所述虚拟现实头戴显示设备，且各第二样本图像中包含所述指定用户的脸部状态；

从所述第一样本图像中识别出第一被遮挡区域，获取所述第一被遮挡区域信息；

根据所述第一被遮挡区域信息，对所述第二样本图像的所述指定目标脸部的与所述第一被遮挡区域对应的区域进行标记；

将所述第二样本图像中标记区域的图像放入第一指定集合中，将该第一指定集合作为深度神经网络训练时的输出集合；将所述第二样本图像中的所述指定目标脸部的未被标记区域的图像放入第二指定集合中，将该第二指定集合作为深度神经网络训练时的输入集合；所述第二指定集合与所述第一指定集合中的图像元素有一一对应的严格的输入输出对应关系；

将所述输入集合和所述输出集合中的每一对具有输入输出对应关系的图像元素输入到预设的深度神经网络中进行训练，确定未遮挡区域图像和生成的与其匹配的遮挡区域图像之间的函数关系，以使所述第二获取单元将所述第一脸部图像数据输入到所述预设的脸

部表情模型,所述预设的脸部表情模型根据输入的所述第一脸部图像数据和所述函数关系输出与其匹配的第二脸部图像数据。

9. 如权利要求7所述的图像处理装置,其特征在于,所述装置还包括:

处理单元,用于从所述实际图像中识别出被所述虚拟现实头戴设备遮挡的非脸部区域;从所述视频流中获取所述实际图像之前的多个第三图像,从所述第三图像中提取背景图像,使用所述背景图像中与被所述虚拟现实头戴设备遮挡的非脸部区域对应的图像数据,对被所述虚拟现实头戴设备遮挡的非脸部区域进行去遮挡处理,或者,从所述实际图像中识别出被所述虚拟现实头戴设备遮挡的非脸部图像数据,将所述非脸部图像数据输入到预设的非脸部模型中,以使所述预设的非脸部模型识别所述非脸部图像数据,输出与被所述非脸部区域匹配的第四图像数据,根据所述第四图像数据对所述非脸部区域进行去遮挡处理。

10. 一种终端设备,其特征在于,所述终端设备包括:如权利要求7-9任一项所述的图像处理装置。

一种图像处理方法、装置和终端设备

技术领域

[0001] 本发明涉及计算机技术领域,特别涉及一种图像处理方法、装置和终端设备。

背景技术

[0002] 虚拟现实技术(Virtual Reality,简称VR)的一个重要应用领域是社交领域。例如,VR视频直播的应用中,主持人侧配置360度摄像头,采集直播地点的全视角场景视频,经由网络共享给接入端的VR头戴显示设备(Head Mounted Device,简称HMD),访客通过佩戴VR HMD体验主持人侧的场景视频,并可以通过转动头部来观看不同视角的场景。该应用的特点是VR视频数据流为单向传输。随着VR社交需求的不断提升,VR社交需要两点之间实现VR视频数据流的双向流动,即社交双方都需要同时配置360度摄像头和VR HMD,同时采集本地全视角视频并发送给对方,由对方从VR HMD 中观看。

[0003] 但是,因为社交双方均佩戴VR HMD,这将导致本地摄像头拍摄到的人脸都会被VR HMD遮挡住眼睛及周围部分。因为眼部周围图像带有非常丰富的表情信息,表情信息的缺失严重影响VR技术在社交领域的应用。所以,急需一种图像处理方案,对被VR HMD遮挡住的眼睛及周围部分进行重建,以保证社交过程中表情信息的完整。

发明内容

[0004] 鉴于上述问题,提出了本发明的一种图像处理方法、装置和终端设备,以便解决或至少部分地解决上述问题。

[0005] 根据本发明的一个方面,提供了一种图像处理方法,该方法包括:

[0006] 从摄像头采集的视频流中获取指定目标的实际图像,其中,指定目标佩戴有虚拟现实头戴设备;

[0007] 从实际图像中识别出指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域,获取与未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据;

[0008] 根据第一脸部图像数据和预设的脸部表情模型,得到与第一脸部图像数据匹配的第二脸部图像数据,第二脸部图像数据与被虚拟现实头戴显示设备遮挡区域相对应;

[0009] 将第一脸部图像数据和第二脸部图像数据相融合,生成合成图像。

[0010] 根据本发明的另一个方面,提供了一种图像处理装置,该装置包括:

[0011] 第一获取单元,用于从摄像头采集的视频流中获取指定目标的实际图像,其中,指定目标佩戴有虚拟现实头戴设备;

[0012] 识别单元,用于从实际图像中识别出指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域,获取与未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据;

[0013] 第二获取单元,用于根据第一脸部图像数据和预设的脸部表情模型,得到与第一脸部图像数据匹配的第二脸部图像数据,第二脸部图像数据与被虚拟现实头戴显示设备遮

挡区域相对应；

[0014] 生成单元,用于将第一脸部图像数据和第二脸部图像数据相融合,生成合成图像。

[0015] 根据本发明的又一个方面,提供了一种终端设备,该终端设备包括:如前所述的图像处理装置。

[0016] 综上所述,本发明技术方案的有益效果是:当获取到戴有虚拟现实头戴设备的指定目标的实际图像后,先从实际图像中识别出指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域,将未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据输入到预设的脸部表情模型中,就可以得到与第一脸部图像数据匹配的第二脸部图像数据;然后将第一脸部图像数据和第二脸部图像数据相融合,生成合成图像。因为第二脸部图像数据与被虚拟现实头戴显示设备遮挡区域相对应,且带有表情信息,所以合成图像则是完整的带有表情信息的图像,相比较使用静态图片来说,合成图像更加逼真、准确,有利于社交双方及时获得对方的表情信息,提高社交质量,保证社交的顺利进行,提升用户体验。

附图说明

[0017] 图1为本发明一个实施例提供的一种图像处理方法的流程示意图；

[0018] 图2为本发明一个实施例提供的一种图像处理装置的功能结构示意图；

[0019] 图3为本发明另一个实施例提供的一种图像处理装置的功能结构示意图；

[0020] 图4为本发明一个实施例提供的一种终端设备的功能结构示意图。

具体实施方式

[0021] 本发明的设计思路是:鉴于使用眼部的静态图片覆盖被VR HMD遮挡的部分仍然会导致表情信息的缺失,且静态图片与脸部其余部分不能很好的融合,会很 unnatural。又考虑到,人脸被虚拟现实头戴显示设备遮挡的眼部及周边部分的图像,与未被虚拟现实头戴显示设备遮挡的脸部图像信息之间有着强相关的关系。本技术方案引入脸部表情模型,通过脸部表情模型得到与未被虚拟现实头戴显示设备遮挡的脸部图像信息匹配的遮挡区域的脸部图像,进而获得具有完整表情信息的合成图像。为使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明实施方式作进一步地详细描述。

[0022] 图1为本发明一个实施例提供的一种图像处理方法的流程示意图。如图 1所示,该方法包括:

[0023] 步骤S110,从摄像头采集的视频流中获取指定目标的实际图像,其中,指定目标佩戴有虚拟现实头戴设备(VR HMD)。

[0024] 在社交应用中,摄像头通过采集指定目标的视频流传输给社交对方。在本实施例中,该指定目标为可以是佩戴VR HMD进行社交的用户,指定目标佩戴着VR HMD,所以实际图像中,指定目标的人脸的眼睛以及眼睛周围部分是被VR HMD遮挡的,无法获取到完整的表情信息,影响社交过程。为了对摄像头采集的图像进行处理,需要从摄像头采集的视频流中获取一指定目标的实际图像。

[0025] 步骤S120,从实际图像中识别出指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域,获取与未被虚拟现实头戴显示设备遮挡区域对

应的第一脸部图像数据。

[0026] 本实施例中,通过图像识别方法识别出实际图像中指定目标的脸部,并识别出脸部未被VR HMD遮挡的区域和被VR HMD遮挡的区域,因为需要通过未被VR HMD遮挡区域,得到与被VR HMD遮挡区域对应的图像数据,所以需要获取与未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据。

[0027] 步骤S130,根据第一脸部图像数据和预设的脸部表情模型,得到与第一脸部图像数据匹配的第二脸部图像数据,第二脸部图像数据与被虚拟现实头戴显示设备遮挡区域相对应。

[0028] 本实施例中,预设的脸部表情模型是通过指定目标样本训练得到的(例如,使用神经网络进行机器学习),在样本训练中可以获得未被VR HMD遮挡的图像数据与被VR HMD遮挡区域的图像数据之间的关系,因此,根据从实际图像中获取的第一脸部图像数据和预设的脸部表情模型,就可以得到与第一脸部图像数据匹配的第二脸部图像数据,即得到与被VR HMD遮挡区域对应的图像数据。

[0029] 针对一个用户来说,只需要进行一次样本训练就可以,但是当用户更换 VR HMD时,因为会存在更换前和更换后的VR HMD的大小不一致的情况,需要进行重新训练,防止根据原预设的脸部表情模型生成的第二脸部图像与第一脸部图像数据不能进行完美的融合。

[0030] 步骤S140,将第一脸部图像数据和第二脸部图像数据相融合,生成合成图像。

[0031] 通过图像融合方法,将第一脸部图像数据和第二脸部图像数据相融合,生成合成图像。因为,第二脸部图像数据是与被VR HMD遮挡区域对应的带有表情信息的图像,因此,合成图像中带有指定目标的完整表情,获得合成图像后,就可以将该合成图像发送社交的对方。

[0032] 因为第二脸部图像数据与被虚拟现实头戴显示设备遮挡区域相对应,且带有表情信息,所以合成图像则是完整的带有表情信息的图像,相比较静态图片来说,合成图像更加逼真、准确,有利于社交双方及时获得对方的表情信息,提高社交质量,保证社交的顺利进行,提升用户体验。

[0033] 在本发明的一个实施例中,步骤S130中的根据第一脸部图像数据和预设的脸部表情模型,得到与第一脸部图像数据匹配的第二脸部图像数据包括:将第一脸部图像数据输入到预设的脸部表情模型中,以使脸部表情模型识别第一脸部图像数据,输出与第一脸部图像数据相匹配的第二脸部图像数据。

[0034] 如上文说明,在预设的脸部表情模型中有未被VR HMD遮挡的图像数据与被VR HMD遮挡区域的图像数据之间的关系,当将第一脸部图像数据输入到预设的脸部表情模型后,脸部表情模型识别第一脸部图像数据,就会输出与第一脸部图像数据相匹配的第二脸部图像数据。也就是说,预设的脸部表情模型会自动分析第一脸部图像数据,然后根据第一脸部图像数据直接生成与第一脸部数据匹配的第二脸部图像数据,有利于提高图像处理的效率,进一步增加用户体验。

[0035] 进一步地,上述的预设的脸部表情模型是通过深度神经网络实现得到的,通过深度神经网络实现得到预设的脸部表情模型包括:

[0036] (1) 获取摄像头在第一场景下采集的指定目标的多个第一样本图像,以及在第二场景下采集的指定目标的多个第二样本图像;其中,在第一场景下,指定目标佩戴有虚拟现

实头戴设备;在第二场景下,指定目标未佩戴虚拟现实头戴显示设备,且各第二样本图像中包含指定用户的脸部状态。

[0037] 在本实施例中,获取多个第一样本图像的目的是为了可以将第二样本图像中与VR HMD遮挡区域对应的部分提取出来。多个第二样本图像中应该包含用户各种表情信息,以便在对实际图像进行处理时,可以匹配到更加准确的第二图像数据。

[0038] (2) 从第一样本图像中识别出第一被遮挡区域,获取第一被遮挡区域信息。

[0039] 如上文说明,为了将第二样本图像中与VR HMD遮挡区域对应的部分提取出来,需要识别出第一样本图像中的第一被遮挡区域,然后获取到第一被遮挡区域信息,例如,区域边界的坐标信息。

[0040] (3) 根据第一被遮挡区域信息,对第二样本图像的指定目标脸部的与第一被遮挡区域对应的区域进行标记。

[0041] (4) 将第二样本图像中标记区域的图像放入第一指定集合中,将该第一指定集合作为深度神经网络训练时的输出集合;将第二样本图像中的指定目标脸部的未被标记区域的图像放入第二指定集合中,将该第二指定集合作为深度神经网络训练时的输入集合,放入第一指定集合和第二指定集合中的图像作为相应集合中的图像元素。其中,第二指定集合与第一指定集合中的图像元素有一一对应的严格的输入输出对应关系。

[0042] (5) 将输入集合和输出集合中的每一对具有输入输出对应关系的图像元素输入到预设的深度神经网络中进行训练,确定未遮挡区域图像和生成的与其匹配的遮挡区域图像之间的函数关系,以使在第一脸部图像数据输入到预设的脸部表情模型时,预设的脸部表情模型根据识别的第一脸部图像数据和函数关系输出与其匹配的第二脸部图像数据。

[0043] 本实施例,设计一个深度神经网络,其类型、层数以及每一层的节点数量,根据图像分辨率和所需生成效果设定。采用深度神经网络的机器学习方法,通过对指定目标的样本图像进行机器学习,获得对应的脸部表情模型。且,本实施例第二指定集合与第一指定集合中的图像元素有一一对应的严格的输入输出对应关系,也就是说,本实施例通过深度神经网络进行有监督式的训练,将具有输入输出对应关系的图像元素输入到深度神经网络中进行训练生成神经网络模型参数,因为输入的图像元素和输出的图像元素有对应关系,通过训练就可以生成未遮挡区域图像和生成的与其匹配的遮挡区域图像之间的函数关系: $output=f(input)$, $input$ 为脸部未遮挡区域的图像, $output$ 则为生成的眼部及周围对应于遮挡区域的脸部图像。

[0044] 可见,本实施例引入深度神经网络的机器学习方法,对指定目标的样本图像进行训练,利用人工智能通过对指定目标的样本图像训练-预测的方式来生成被VR HMD遮挡区域的图像数据,可以使得合成图像与指定目标更加匹配,生成的合成图像更加自然,增强用户体验。

[0045] 具体地,在预设的深度神经网络训练过程中,预设的深度神经网络训练的损失函数是输出集合中的图像和生成的与输入集合中的图像相匹配的图像之间的均方差。

[0046] 在实际应用中,VR HMD比指定目标的脸部要大,图像中除了指定目标的脸部区域的部分,VR HMD还会遮挡一部分非脸部区域,如果仅对脸部进行图像处理,生成的合成图像与真实效果的差距较大,需要对被VR HMD遮挡的非脸部图像进行去遮挡处理,可以通过下述的方法进行:

[0047] (1) 在本发明的一个实施例中,图1所示的方法还包括:从实际图像中识别出被虚拟现实头戴设备遮挡的非脸部区域;从视频流中获取实际图像之前的多个第三图像,从第三图像中提取背景图像,使用背景图像中与被虚拟现实头戴设备遮挡的非脸部区域对应的图像数据,对被虚拟现实头戴设备遮挡的非脸部区域进行去遮挡处理。

[0048] 这里第三图像的个数不具体限定。因为摄像头采集视频流是与环境的位置是相对固定的,可以根据实际图像之前的多个图像帧中的背景图像信息进行去遮挡处理。

[0049] (2) 在本发明的另一个实施例中,图1所示的方法还包括:从实际图像中识别出被虚拟现实头戴设备遮挡的非脸部图像数据,将非脸部图像数据输入到预设的非脸部模型中,以使预设的非脸部模型识别非脸部图像数据,输出与被虚拟现实头戴设备遮挡的非脸部区域匹配的第四图像数据,根据第四图像数据对被虚拟现实头戴设备遮挡的非脸部区域进行去遮挡处理。

[0050] 本实施例中预设的非脸部模型可以通过无监督训练的神经网络生成。上述的去遮挡处理可以采用图像融合方法,将获取的与被VR HMD遮挡的非脸部区域对应的图像数据或者第四图像数据与实际图像中未被VR HMD遮挡的图像数据进行融合。

[0051] 通过上述的(1)和(2)对被虚拟现实头戴设备遮挡的非脸部区域,避免第一脸部图像数据和第二脸部图像数据融合后,与非脸部区域的衔接处过于明显,保证生成的合成图像更加真实、完整,而非仅仅体现指定目标的表情信息,整个合成图像更具有观赏性,增强用户体验。

[0052] 在一个实施例中,该图像处理方法在实际应用中,生成合成图像是将第一脸部图像数据、第二脸部图像数据、非人脸部分中未被VR HMD遮挡的图像数据,以及获取的与被VR HMD遮挡的非脸部区域对应的图像数据或者第四图像数据进行融合,以生成完整的合成图像。

[0053] 例如,本实施例中被VR HMD遮挡的非脸部图像数据可以是指定目标的头发或耳朵等区域,通过上述的(1)或(2)就可以将被遮挡的头发或耳朵展现出来,使得生成的合成图像更加逼真。

[0054] 图3为本发明一个实施例提供的一种图像处理装置的功能结构示意图。如图3所示,该图像处理装置200包括:

[0055] 第一获取单元210,用于从摄像头采集的视频流中获取指定目标的实际图像,其中,指定目标佩戴有虚拟现实头戴设备。

[0056] 识别单元220,用于从实际图像中识别出指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域,获取与未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据。

[0057] 第二获取单元230,用于根据第一脸部图像数据和预设的脸部表情模型,得到与第一脸部图像数据匹配的第二脸部图像数据,第二脸部图像数据与被虚拟现实头戴显示设备遮挡区域相对应。

[0058] 生成单元240,用于将第一脸部图像数据和第二脸部图像数据相融合,生成合成图像。

[0059] 在本发明的一个实施例中,第二获取单元230,用于将第一脸部图像数据输入到预设的脸部表情模型中,以使脸部表情模型识别第一脸部图像数据,输出与第一脸部图像数

据相匹配的第二脸部图像数据。

[0060] 在本发明的一个实施例中,第二获取单元230还包括:

[0061] 训练模块,用于通过神经网络实现得到预设的脸部表情模型,具体用于:获取摄像头在第一场景下采集的指定目标的多个第一样本图像,以及在第二场景下采集的指定目标的多个第二样本图像;其中,在第一场景下,指定目标佩戴有虚拟现实头戴设备;在第二场景下,指定目标未佩戴虚拟现实头戴显示设备,且各第二样本图像中包含指定用户的脸部状态;从第一样本图像中识别出第一被遮挡区域,获取第一被遮挡区域信息;根据第一被遮挡区域信息,对第二样本图像的指定目标脸部的与第一被遮挡区域对应的区域进行标记;将第二样本图像中标记区域的图像放入第一指定集合中,将该第一指定集合作为神经网络训练时的输出集合;将第二样本图像中的指定目标脸部的未被标记区域的图像放入第二指定集合中,将该第二指定集合作为神经网络训练时的输入集合;第二指定集合与第一指定集合中的图像元素有一一对应的严格的输入输出对应关系;将输入集合和输出集合中的每一具有对输入输出对应关系的图像元素输入到预设的神经网络中进行训练,确定未遮挡区域图像和生成的与其匹配的遮挡区域图像之间的函数关系,以使第二获取单元将第一脸部图像数据输入到预设的脸部表情模型,预设的脸部表情模型根据输入的第一脸部图像数据和函数关系输出与其匹配的第二脸部图像数据。

[0062] 进一步地,在预设的神经网络训练过程中,预设的神经网络训练的损失函数是输出集合中的图像和生成的与输入集合中的图像相匹配的图像之间的均方差。

[0063] 在本发明的一个实施例中,图2所示的装置还包括:

[0064] 处理单元,用于从实际图像中识别出被虚拟现实头戴设备遮挡的非脸部区域;从视频流中获取实际图像之前的多个第三图像,从第三图像中提取背景图像,使用背景图像中与被虚拟现实头戴设备遮挡的非脸部区域对应的图像数据,对被虚拟现实头戴设备遮挡的非脸部区域进行去遮挡处理,或者,从实际图像中识别出被所述虚拟现实头戴设备遮挡的非脸部图像数据,将非脸部图像数据输入到预设的非脸部模型中,以使预设的非脸部模型识别非脸部图像数据,输出与被非脸部区域匹配的第四图像数据,根据第四图像数据对非脸部区域进行去遮挡处理。

[0065] 图3为本发明另一个实施例提供的一种图像处理装置的结构示意图。如图3所示,图像处理装置300包括存储器310和处理器320,存储器310和处理器320之间通过内部总线330通讯连接,存储器310存储有能够被处理器320执行的图像处理的计算机程序311,该图像处理的计算机程序311被处理器320执行时能够实现图1中所示的方法步骤。

[0066] 在不同的实施例中,存储器310可以是内存或者非易失性存储器。其中非易失性存储器可以是:存储驱动器(如硬盘驱动器)、固态硬盘、任何类型的存储盘(如光盘、DVD等),或者类似的存储介质,或者它们的组合。内存可以是:RAM(Random Access Memory,随机存取存储器)、易失存储器、非易失性存储器、闪存。进一步,非易失性存储器和内存作为机器可读存储介质,其上可存储由处理器320执行的图像处理的计算机程序311。

[0067] 图4为本发明一个实施例提供的一种终端设备的功能结构示意图。如图4所示,该终端设备400包括:如图3或图4所示的图像处理装置410。

[0068] 在本发明的一个实施例中,该终端设备410是虚拟现实头戴显示设备。或者,该终端设备410是在社交过程中与虚拟现实头戴显示设备进行连接的计算机或服务器,合成图

像可以通过计算机或者服务器发送给社交的对方。

[0069] 需要说明的是,图2、图3所示的装置和图4所示的终端设备的各实施例与图1所示的方法的各实施例对应相同,上文已有详细说明,在此不再赘述。

[0070] 综上所述,本发明技术方案的有益效果是:当获取到戴有虚拟现实头戴设备的指定目标的实际图像后,先从实际图像中识别出指定目标脸部的未被虚拟现实头戴显示设备遮挡区域和被虚拟现实头戴显示设备遮挡区域,将未被虚拟现实头戴显示设备遮挡区域对应的第一脸部图像数据输入到预设的脸部表情模型中,就可以得到与第一脸部图像数据匹配的第二脸部图像数据;然后将第一脸部图像数据和第二脸部图像数据相融合,生成合成图像。因为第二脸部图像数据与被虚拟现实头戴显示设备遮挡区域相对应,且带有表情信息,所以合成图像则是完整的带有表情信息的图像,有利于社交双方及时获得对方的表情信息,提高社交质量,保证社交的顺利进行,提升用户体验。

[0071] 以上所述,仅为本发明的具体实施方式,在本发明的上述教导下,本领域技术人员可以在上述实施例的基础上进行其他的改进或变形。本领域技术人员应该明白,上述的具体描述只是更好的解释本发明的目的,本发明的保护范围应以权利要求的保护范围为准。

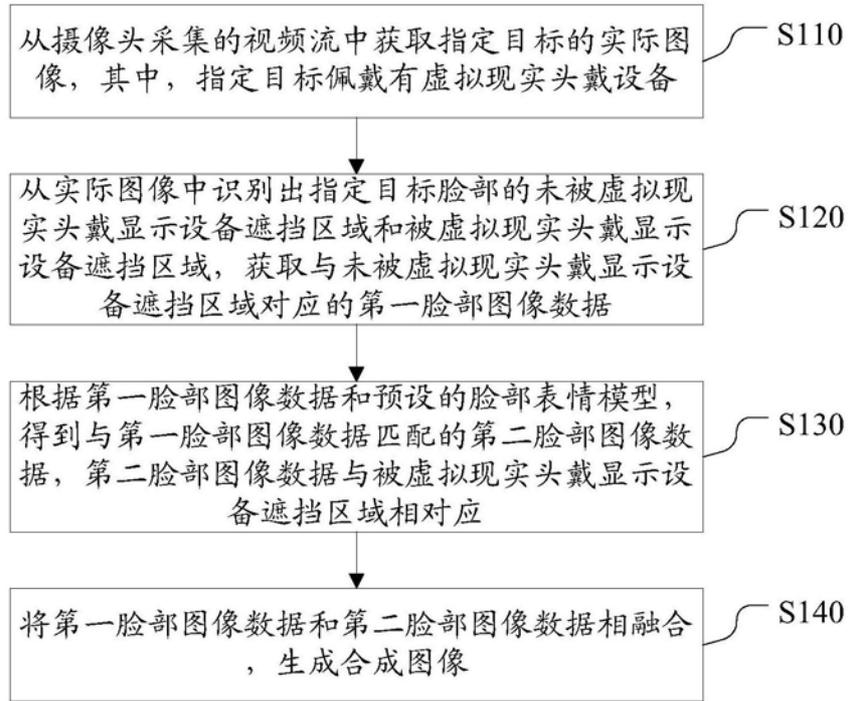


图1

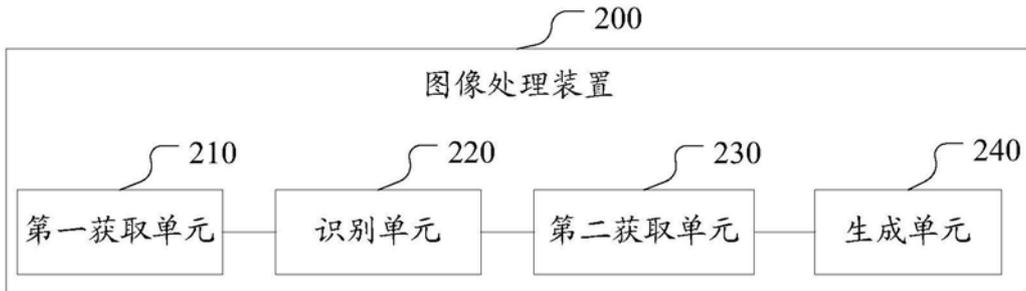


图2

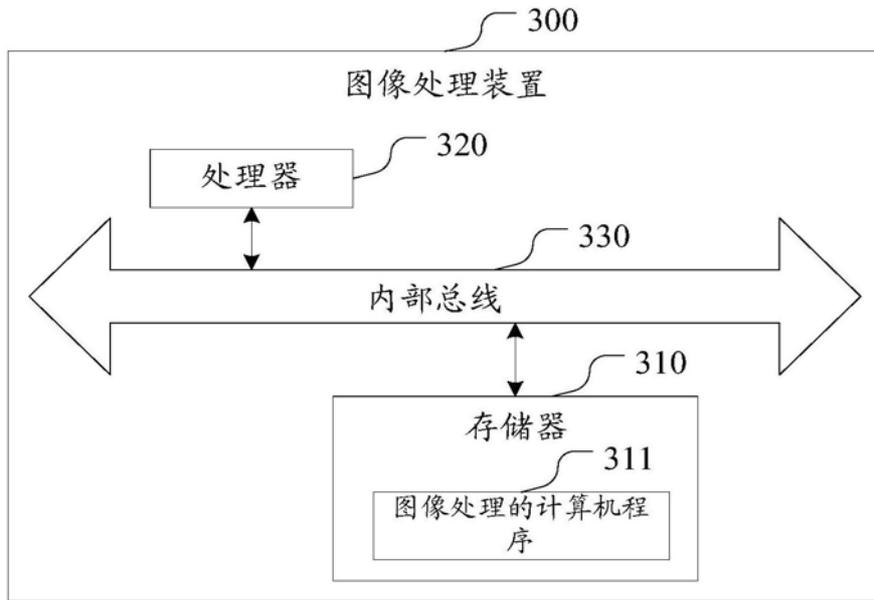


图3



图4